

# Sequence Alignment

GenBank Accession Number AB020673

AB020673

LOCUS AB020673 6846 bp mRNA linear

DEFINITION Homo sapiens mRNA for KIAA0866 protein, partial cds.

ACCESSION AB020673

VERSION AB020673.2 GI:27529743

KEYWORDS

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1

AUTHORS Nagase, T., Kikuno, R., Yamakawa, H. and Ohara, O.

TITLE The nucleotide sequence of a long cDNA clone isolated from human

JOURNAL Published Only in Database (1999)

REFERENCE 2 (bases 1 to 6846)

AUTHORS Nagase, T., Kikuno, R., Yamakawa, H. and Ohara, O.

TITLE Direct Submission

JOURNAL Submitted (02-DEC-1998) Osamu Ohara, Kazusa DNA Research Institute,  
Department of Human Gene Research; 2-6-7 Kazusa-kamatari, Kisarazu,  
Chiba 292-0818, Japan (E-mail: cdnainfo@kazusa.or.jp,  
URL: http://www.kazusa.or.jp/huge, Tel: 81-438-52-3913,  
Fax: 81-438-52-3914)

COMMENT On Jan 7, 2003 this sequence version replaced gi:4240220.

FEATURES

source

Location/Qualifiers

1..6846

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="hk00546s1"

/tissue\_type="brain"

/clone\_lib="pBluescriptII SK plus"

gene

1..6846

/gene="KIAA0866"

CDS

<40..5994

/gene="KIAA0866"

/note="This sequence was replaced that of hk06733 cDNA as  
a representative cDNA sequence for KIAA0866."

/codon\_start=1

/product="KIAA0866 protein"

/protein\_id="BAA74889.2"

/db\_xref="GI:27529744"

/translation="SPQLEIWDQQGTMAQKGQLSDDEKFLFVDKNFINSPVAQADWAA  
KRLVWVPSEKQGFEEASIKKEEGDEVVVELVENGKKVTVGKDDIQMNPFPKSKVEDM  
AELTCLNEASVLHNLRLERYFSGLIYTYSGLFVNVNPKHLPIYSEKIVDMYKGGKRH  
EMPPHIYAIADTAYRSMQLDREDQSILCTGESGAGKTENTKKVIQYLAVVASSHKGKK  
DTSITGELEKQLLQANPILEAFNAKTVKNDNSSRFGKFI RINFDVTGYIVGANIETY  
LLEKSRAIRQARDERTFHIYYMIAGAKEKMRSDLLLEGFNNYTFLSNGFVPIPAQD  
DEMFOETVEAMAIMGFSEEEQLSILKVVSSVLQGNIVFKKERNTDQASMPDNTAAQK  
VCHLMGINVTDFTRSILTPRIKVGDRDVVQKAQTKEQADFVEALAKATYERLFRWILT  
RVNKALDKTHRQGSFGLIDIAGFEIFEVNSFEQLCINYNTEKLQQLFNHTMFILQ  
EEYQREGIEWNFIDFGLDLQPCIELIERPNNPPGVLLALLDEECWFPKATDKSFVEKLC  
TEQGSHPKFKPKQLKDKTEFSIIHYAGKVDYNASAWLTKNMDPLNDNVTSLLNASSD  
KFVADLWKDVDRIVGLDQMAKMTESLPSASKTKKGMFRTVGQLYKEQLGKLMTTLRN  
TTPNFVRCIIPNHEKRSGLDAFLVLEQLRCNGVLEGIRICRQGFPPNRIVFQEFQRQY  
EILANAIPKGFMDGKQACILMIKALELDPNLYRIGQSKIFFRTGVLAHLEERDLKI  
TDVIMAFQAMCRGYLARKAFKQQQLTAMKVIQRNCAAYLKLNRNQQWRLFTKVKPL  
LQVTRQEEEMQAKEDQLTKERQQAENELKELEQKHSQLETEKNLLQEQLQAETEL  
YAAEAEEMRVRLAAKQLEELIHEMEARLEEEEDRQQLQAEKKAQAQMLDLEEQLE  
BEEAARQKLQLEKVTAEAKIKKLEDEILVMDQNNKLSKERKLEERISDLTTNLAE  
EEKAKNLTILKNKHESMISELEVLRLKKEEKSQLEKLRKLEGDASDFHEQIADLQA  
QIAELKMLAKKEELQAALARLDDEIAQKNNAKKIRELEGHISDLQEDLDSERAAR  
NKAQKQKRLDGEELAKTELEDTLSTATQQLRAKREQEVTVLKALDEETRSHEA  
VQEMRQKHAQAVEELTEQLEQFKRANLNDKNQTLKENADLAGELRVLGQAKQEV  
EHKKKLEAQVQELQSKSDGERARAELNDKVHKLQNEVESVTGMLNEAEGKAIKIAK  
DVASLSSQLQDTQELLQEETRQKLNVTSLRQLEEEERNSLQDQLDEMEAKQNLERHI  
STLNLQLSDSKKQLQDFASTVEALEEGKKRFQKEIENLTQQYEEKAAAYDKLEKTKNR  
LQQLDDLVVDLNDQRQLVSNLEKKQRKFDQLLAEEKNISSKYADERDRABAEAREKE  
TKALSLARALEEAEKEELERTNKMLKAEMEDLVSSKDDVGKNVHELEKSKRALETQ  
MEEMKTQLELEDELQATEDAKLRLEVNMQALKGQFERDLQARDEQNEEKRRQLQRQL  
HEYETELEDERKQALAAAANKKLEGDLKDELQADSAIKGEEAIKQLRKLQAQMKD

FQRELEDARASRDEIFATAKENEKKAKSLEADLMQLQEDLAAAERARKQADLEKEELA  
 BELASSLSGRNALQDEKRRLEARIAQLEEELEEEQGNMEAMSDRVRKATQQAQELSNE  
 LATERSTAQKNESARQQLERQNKELRSKLHEMEGAVKSKFKSTIAALEAKIAQLEEQV  
 EQEAREKQAATKSLKQKDKKLKEILLQVEDERKMAEQYKEQAEKGNARVKQLKRQLEE  
 AEEESQRINANRRKLQRELDEATESNEAMGREVNALKSKLRRGNETSFVPSRRSGGRR  
 VIENADGSEETDTRDADFNGTKASE"

ORIGIN

Alignment Scores:

Pred. No.:	0	Length:	6846
Score:	9602.50	Matches:	1929
Percent Similarity:	99.64%	Conservative:	0
Best Local Similarity:	99.64%	Mismatches:	0
Query Match:	98.93%	Indels:	7
DB:	9	Gaps:	1

US-09-927-597-2 (1-1945) x AB020673 (1-6846)

Qy	1	MetAlaGlnLysGlyGlnLeuSerAspAspGluLysPheLeuPheValAspLysAsnPhe	20
Db	76	ATGCGCGCAGAAGGCCAACTCAGTGACGATGAGAAGTTCCTCTTGTGGACAAAACCTC	135
Qy	21	IleAsnSerProValAlaGlnAlaAspTrpAlaAlaLysArgLeuValTrpValProSer	40
Db	136	ATCAACAGCCCAGTGGCCAGGCTGACTGGGCCGCCAAGAGACTCGTCTGGGTCCCCTCG	195
Qy	41	GluLysGlnGlyPheGluAlaAlaSerIleLysGluGluLysGlyAspGluValValVal	60
Db	196	GAGAAGCAGGGCTTCGAGGCAGCCAGCATTAAAGGAGGAGAAGGGGGATGAGGTGGTTGTG	255
Qy	61	GluLeuValGluAsnGlyLysLysValThrValGlyLysAspAspIleGlnLysMetAsn	80
Db	256	GAGCTGGTGGAGAATGGCAAGAAGGTCACGGTTGGGAAAGATGACATCCAGAAGATGAAC	315
Qy	81	ProProLysPheSerLysValGluAspMetAlaGluLeuThrCysLeuAsnGluAlaSer	100
Db	316	CCACCCAAGTTCTCCAAGGTGGAGGACATGGCGGAGCTGACGTGCCTCAACGAAGCCTCC	375
Qy	101	ValLeuHisAsnLeuArgGluArgTyrPheSerGlyLeuIleTyrThrTyrSerGlyLeu	120
Db	376	GTGCTACACAACCTGAGGGAGCGGTACTTCTCAGGGCTAATATATACGTACTCTGGCCTC	435
Qy	121	PheCysValValValAsnProTyrLysHisLeuProIleTyrSerGluLysIleValAsp	140
Db	436	TTCTGCGTGGTGGTCAACCCCTATAAACACCTGCCATCTACTCGGAGAAGATCGTCGAC	495
Qy	141	MetTyrLysGlyLysLysArgHisGluMetProProHisIleTyrAlaIleAlaAspThr	160
Db	496	ATGTACAAGGGCAAGAAGAGGCACGAGATGCCGCTCACATCTACGCCATCGCAGACACG	555
Qy	161	AlaTyrArgSerMetLeuGlnAspArgGluAspGlnSerIleLeuCysThrGlyGluSer	180
Db	556	GCCTACCGGAGCATGCTTCAAGATCGGGAGGACCAGTCCATTCTATGCACAGGCGAGTCT	615
Qy	181	GlyAlaGlyLysThrGluAsnThrLysLysValIleGlnTyrLeuAlaValValAlaSer	200
Db	616	GGAGCCGGGAAAACCGAAAACACCAAGAAGGTCATTCAGTACCTGGCCGTGGTGGCCTCC	675
Qy	201	SerHisLysGlyLysLysAspThrSerIleThrGlnGlyProSerPheAlaTyrGlyGlu	220
Db	676	TCCACAAGGGCAAGAAGACACAAGTATCAG-----GGAGAG	714
Qy	221	LeuGluLysGlnLeuLeuGlnAlaAsnProIleLeuGluAlaPheGlyAsnAlaLysThr	240
Db	715	CTGGAAAAGCAGCTTCTACAAGCAAACCCGATTCTGGAGGCTTTCGGCAACGCCAAAACA	774
Qy	241	ValLysAsnAspAsnSerSerArgPheGlyLysPheIleArgIleAsnPheAspValThr	260
Db	775	GTGAGAAGACGACAACTCCTCAGGATTCTGGCAAATTCATCCGCATCAACTTCAGGTCACG	834
Qy	261	GlyTyrIleValGlyAlaAsnIleGluThrTyrLeuLeuGluLysSerArgAlaIleArg	280

Db	835		GGTTACATCGTGGGAGCCAACATTGAGACCTATCTGCTAGAAAAATCACGGGCAATTTCGC	894
Qy	281		GlnAlaArgAspGluArgThrPheHisIlePheTyrTyrMetIleAlaGlyAlaLysGlu	300
Db	895		CAAGCCAGAGACGAGAGGACATTCCACATCTTTACTACATGATTGCTGGAGCCAAGGAG	954
Qy	301		LysMetArgSerAspLeuLeuLeuGluGlyPheAsnAsnTyrThrPheLeuSerAsnGly	320
Db	955		AAGATGAGAAGTGACTTGCTTTTGGAGGGCTTCAACAACCTACACCTTCTCTCCAATGGC	1014
Qy	321		PheValProIleProAlaAlaGlnAspAspGluMetPheGlnGluThrValGluAlaMet	340
Db	1015		TTTGTGCCCATCCAGCAGCCAGGATGATGAGATGTTCCAGGAAACCGTGGAGGCCATG	1074
Qy	341		AlaIleMetGlyPheSerGluGluGluGlnLeuSerIleLeuLysValValSerSerVal	360
Db	1075		GCAATCATGGGTTTCAGCGAGGAGGAGCAGCTATCCATATTGAAGGTGGTATCATCGGTC	1134
Qy	361		LeuGlnLeuGlyAsnIleValPheLysLysGluArgAsnThrAspGlnAlaSerMetPro	380
Db	1135		CTGCAGCTTGGAAATATCGTCTTCAAGAAGGAAAGAAACACAGACCAGGCGTCCATGCCA	1194
Qy	381		AspAsnThrAlaAlaGlnLysValCysHisLeuMetGlyIleAsnValThrAspPheThr	400
Db	1195		GATAACACAGCTGCTCAGAAAGTTGCCACCTCATGGGAATTAATGTGACAGATTTACC	1254
Qy	401		ArgSerIleLeuThrProArgIleLysValGlyArgAspValValGlnLysAlaGlnThr	420
Db	1255		AGATCCATCCTCACTCTCGTATCAAGGTTGGGCGAGATGTGGTACAGAAAGCTCAGACA	1314
Qy	421		LysGluGlnAlaAspPheAlaValGluAlaLeuAlaLysAlaThrTyrGluArgLeuPhe	440
Db	1315		AAAGAACAGGCTGACTTTGCTGTAGAGGCTTTGGCCAAGGCAACATATGAGCGCCTTTTC	1374
Qy	441		ArgTrpIleLeuThrArgValAsnLysAlaLeuAspLysThrHisArgGlnGlyAlaSer	460
Db	1375		CGCTGGATACTACCCGCGTGAACAAAGCCCTGGACAAGACCCATCGGCAAGGGGCTTCC	1434
Qy	461		PheLeuGlyIleLeuAspIleAlaGlyPheGluIlePheGluValAsnSerPheGluGln	480
Db	1435		TTCCTGGGGATCCTGGATATAGCTGGATTGAGATCTTTGAGGTGAACCTCTTCGAGCAG	1494
Qy	481		LeuCysIleAsnTyrThrAsnGluLysLeuGlnGlnLeuPheAsnHisThrMetPheIle	500
Db	1495		CTGTGCATCAACTACACCAACGAGAAGCTGCAGCAGCTCTTCAACCACCATGTTTCATC	1554
Qy	501		LeuGluGlnGluGluTyrGlnArgGluGlyIleGluTrpAsnPheIleAspPheGlyLeu	520
Db	1555		CTGGAGCAGGAGGAGTACCAGCGCGAGGGCATCGAGTGAACCTTCATCGACTTTGGGCTG	1614
Qy	521		AspLeuGlnProCysIleGluLeuIleGluArgProAsnAsnProProGlyValLeuAla	540
Db	1615		GACCTACAGCCCTGCATCGAGCTCATCGAGCGACCGAACAACCTCCAGGTGTGCTGGCC	1674
Qy	541		LeuLeuAspGluGluCysTrpPheProLysAlaThrAspLysSerPheValGluLysLeu	560
Db	1675		CTGCTGGACGAGGAATGCTGGTTCCCCAAAGCCACGGACAAGTCTTTCGTGGAGAAGCTG	1734
Qy	561		CysThrGluGlnGlySerHisProLysPheGlnLysProLysGlnLeuLysAspLysThr	580
Db	1735		TGCACGGAGCAGGGCAGCCACCCCAAGTTCCAGAAGCCCAAGCAGCTCAAGGACAAGACT	1794
Qy	581		GluPheSerIleIleHisTyrAlaGlyLysValAspTyrAsnAlaSerAlaTrpLeuThr	600
Db	1795		GAGTTCTCCATCATCCATTATGCTGGGAAGGTGGACTATAATGCGAGTGCCTGGCTGACC	1854
Qy	601		LysAsnMetAspProLeuAsnAspAsnValThrSerLeuLeuAsnAlaSerSerAspLys	620
Db	1855		AAGAATATGGACCCGCTGAATGACAACGTGACTTCCTGCTCAATGCCTCCTCCGACAAG	1914

Qy	621	PheValAlaAspLeuTrpLysAspValAspArgIleValGlyLeuAspGlnMetAlaLys	640
Db	1915	TTTGTGGCCGACCTGTGGAAGGACGTGGACCGCATCGTGGGCCTGGACCAGATGGCCAAG	1974
Qy	641	MetThrGluSerSerLeuProSerAlaSerLysThrLysLysGlyMetPheArgThrVal	660
Db	1975	ATGACGGAGAGCTCGCTGCCAGCGCTCCAAGACCAAGAAGGGCATGTTCCGCACAGTG	2034
Qy	661	GlyGlnLeuTyrLysGluGlnLeuGlyLysLeuMetThrThrLeuArgAsnThrThrPro	680
Db	2035	GGGCAGCTGTACAAGGAGCAGCTGGGCAAGCTGATGACCACGCTACGCAACACCACGCCC	2094
Qy	681	AsnPheValArgCysIleIleProAsnHisGluLysArgSerGlyLysLeuAspAlaPhe	700
Db	2095	AACTTCGTGCGCTGCATCATCCCCAACACGAGAAGAGGTCGGCAAGCTGGATGCGTTC	2154
Qy	701	LeuValLeuGluGlnLeuArgCysAsnGlyValLeuGluGlyIleArgIleCysArgGln	720
Db	2155	CTGGTGCTGGAGCAGCTGCGGTGCAATGGGGTGCTGGAAGGCATTGCGATCTGCCGGCAG	2214
Qy	721	GlyPheProAsnArgIleValPheGlnGluPheArgGlnArgTyrGluIleLeuAlaAla	740
Db	2215	GGCTTCCCCAACCGGATCGTCTTCCAGGAGTTCGCCAACGCTACGAGATCCTGGCGGCG	2274
Qy	741	AsnAlaIleProLysGlyPheMetAspGlyLysGlnAlaCysIleLeuMetIleLysAla	760
Db	2275	AATGCCATCCCCAAGGCTTCATGGACGGGAAGCAGGCCTGCATTCTCATGATCAAAGCC	2334
Qy	761	LeuGluLeuAspProAsnLeuTyrArgIleGlyGlnSerLysIlePhePheArgThrGly	780
Db	2335	CTGGAAGTGGACCCCACTTATACAGGATAGGGCAGAGCAAAATCTTCTCCGAAGTGGC	2394
Qy	781	ValLeuAlaHisLeuGluGluGluArgAspLeuLysIleThrAspValIleMetAlaPhe	800
Db	2395	GTCCTGGCCCACTAGAGGAGGAGCGAGATTTGAAGATCACCGATGTCATCATGGCCTTC	2454
Qy	801	GlnAlaMetCysArgGlyTyrLeuAlaArgLysAlaPheAlaLysArgGlnGlnGlnLeu	820
Db	2455	CAGGCGATGTGTCTGGCTACTTGGCCAGAAAGGCTTTTGCCAAGAGGCAGCAGCAGCTG	2514
Qy	821	ThrAlaMetLysValIleGlnArgAsnCysAlaAlaTyrLeuLysLeuArgAsnTrpGln	840
Db	2515	ACCGCCATGAAGGTGATTGAGAGAACTGCGCCGCCTACCTCAAGCTGCGGAAGTGGCAG	2574
Qy	841	TrpTrpArgLeuPheThrLysValLysProLeuLeuGlnValThrArgGlnGluGluGlu	860
Db	2575	TGGTGGAGGCTTTTCACCAAAGTGAAGCCACTGCTGCAGGTGACACGGCAGGAGGAGGAG	2634
Qy	861	MetGlnAlaLysGluAspGluLeuGlnLysThrLysGluArgGlnGlnLysAlaGluAsn	880
Db	2635	ATGCAGGCCAAGGAGGATGAAGTGCAGAGACCAAGGAGCGGCAGCAGAGAAGGCAGAGAAT	2694
Qy	881	GluLeuLysGluLeuGluGlnLysHisSerGlnLeuThrGluGluLysAsnLeuLeuGln	900
Db	2695	GAGCTTAAGGAGCTGGAACAGAAGCACTCGCAGCTGACCGAGGAGAAGAACCTGCTACAG	2754
Qy	901	GluGlnLeuGlnAlaGluThrGluLeuTyrAlaGluAlaGluGluMetArgValArgLeu	920
Db	2755	GAACAGCTGCAGGCAGAGACAGAGCTGTATGCAGAGGCTGAGGAGATGCGGGTGCGGCTG	2814
Qy	921	AlaAlaLysLysGlnGluLeuGluGluIleLeuHisGluMetGluAlaArgLeuGluGlu	940
Db	2815	GCGGCCAAGAAGCAGGAGCTGGAGGAGATACTGCATGAGATGGAGGCCCGCTGGAGGAG	2874
Qy	941	GluGluAspArgGlyGlnGlnLeuGlnAlaGluArgLysLysMetAlaGlnGlnMetLeu	960
Db	2875	GAGGAAGACAGGGGCCAGCAGCTACAGGCTGAAAGGAAGAAGATGGCCCAGCAGATGCTG	2934
Qy	961	AspLeuGluGluGlnLeuGluGluGluGluAlaAlaArgGlnLysLeuGlnLeuGluLys	980
Db	2935	GACCTTGAAGAACAGCTGGAGGAGGAGGAAGCTGCCAGGCAGAAGCTGCAACTTGAGAAG	2994

Qy	981	ValThrAlaGluAlaLysIleLysLysLeuGluAspGluIleLeuValMetAspAspGln	1000
Db	2995	GTCACGGCTGAGGCCAAGATCAAGAACTGGAGGATGAGATCCTGGTCATGGATGATCAG	3054
Qy	1001	AsnAsnLysLeuSerLysGluArgLysLeuLeuGluGluArgIleSerAspLeuThrThr	1020
Db	3055	AACAATAAACTATCAAAAGAACGAAACTCCTTGAGGAGAGGATTAGTGACTTAACGACA	3114
Qy	1021	AsnLeuAlaGluGluGluGluLysAlaLysAsnLeuThrLysLeuLysAsnLysHisGlu	1040
Db	3115	AATCTTGCAGAAGAGGAAGAAAAGGCCAAGAATCTTACCAAGCTGAAAAACAAGCATGAA	3174
Qy	1041	SerMetIleSerGluLeuGluValArgLeuLysLysGluGluLysSerArgGlnGluLeu	1060
Db	3175	TCTATGATTTTCAGAACTGGAAGTGCGGCTAAAGAAGGAAGAGAAGAGCCGACAGGAGCTG	3234
Qy	1061	GluLysLeuLysArgLysLeuGluGlyAspAlaSerAspPheHisGluGlnIleAlaAsp	1080
Db	3235	GAGAAGCTGAAACGGGAAGCTGGAGGGTGATGCCAGCGACTTCCACGAGCAGATCGCTGAC	3294
Qy	1081	LeuGlnAlaGlnIleAlaGluLeuLysMetGlnLeuAlaLysLysGluGluGluLeuGln	1100
Db	3295	CTCCAGGCGCAGATCGCAGAGCTCAAGATGCAGCTGGCCAAGAAGGAGGAGGAGCTGCAG	3354
Qy	1101	AlaAlaLeuAlaArgLeuAspAspGluIleAlaGlnLysAsnAsnAlaLeuLysLysIle	1120
Db	3355	GCGGCCCTGGCCAGGCTTGACGATGAAATCGCTCAGAAGAACAATGCCCTGAAGAAGATC	3414
Qy	1121	ArgGluLeuGluGlyHisIleSerAspLeuGlnGluAspLeuAspSerGluArgAlaAla	1140
Db	3415	CGGGAGCTGGAGGGCCACATCTCAGACCTCCAGGAGGACCTGGACTCAGAGCGGGCCGCC	3474
Qy	1141	ArgAsnLysAlaGluLysGlnLysArgAspLeuGlyGluGluLeuGluAlaLeuLysThr	1160
Db	3475	AGGAACAAGGCTGAAAAGCAGAAGCGAGACCTCGGCGAGGAGCTGGAGGCCCTAAAGACA	3534
Qy	1161	GluLeuGluAspThrLeuAspSerThrAlaThrGlnGlnGluLeuArgAlaLysArgGlu	1180
Db	3535	GAGCTGGAAGACACACTGGACAGCACAGCCACTCAGCAGGAGCTCAGGGCCAAGAGGGAG	3594
Qy	1181	GlnGluValThrValLeuLysLysAlaLeuAspGluGluThrArgSerHisGluAlaGln	1200
Db	3595	CAGGAGGTGACGGTGCTGAAGAAGGCCCTGCATGAAGAGACGCGGTCCCATGAGGCTCAG	3654
Qy	1201	ValGlnGluMetArgGlnLysHisAlaGlnAlaValGluGluLeuThrGluGlnLeuGlu	1220
Db	3655	GTCCAGGAGATGAGGCAGAAACACGCACAGGCGGTGGAGGAGCTCACAGAGCAGCTTGAG	3714
Qy	1221	GlnPheLysArgAlaLysAlaAsnLeuAspLysAsnLysGlnThrLeuGluLysGluAsn	1240
Db	3715	CAGTTCAAGAGGGCCAAGGCCAACCTAGACAAGAATAAGCAGACGCTGGAGAAAGAGAAC	3774
Qy	1241	AlaAspLeuAlaGlyGluLeuArgValLeuGlyGlnAlaLysGlnGluValGluHisLys	1260
Db	3775	GCAGACCTGGCCGGGAGCTGCGGGTCTGGGCCAGGCCAAGCAGGAGGTGGAACATAAG	3834
Qy	1261	LysLysLysLeuGluAlaGlnValGlnGluLeuGlnSerLysCysSerAspGlyGluArg	1280
Db	3835	AAGAAGAAGCTGGAGGCGCAGGTGCAGGAGCTGCAGTCCAAGTGCAGCGATGGGGAGCGG	3894
Qy	1281	AlaArgAlaGluLeuAsnAspLysValHisLysLeuGlnAsnGluValGluSerValThr	1300
Db	3895	GCCCCGGCGGAGCTCAATGACAAAGTCCACAAGCTGCAGAATGAAGTTGAGAGCGTCACA	3954
Qy	1301	GlyMetLeuAsnGluAlaGluGlyLysAlaIleLysLeuAlaLysAspValAlaSerLeu	1320
Db	3955	GGGATGCTTAACGAGGCCGAGGGGAAGGCCATTAAGCTGGCCAAGGACGTGGCGTCCCTC	4014
Qy	1321	SerSerGlnLeuGlnAspThrGlnGluLeuLeuGlnGluGluThrArgGlnLysLeuAsn	1340

Db 4015 AGTTCACAGCTCCAGGACACCCAGGAGCTGCTTCAAGAAGAAACCCGGCAGAAGCTCAAC 4074  
 Qy 1341 ValSerThrLysLeuArgGlnLeuGluGluGluArgAsnSerLeuGlnAspGlnLeuAsp 1360  
 |||||  
 Db 4075 GTGTCTACGAAGCTGCGCCAGCTGGAGGAGGAGCGGAACAGCCTGCAAGACCAGCTGGAC 4134  
 Qy 1361 GluGluMetGluAlaLysGlnAsnLeuGluArgHisIleSerThrLeuAsnIleGlnLeu 1380  
 |||||  
 Db 4135 GAGGAGATGGAGGCCAAGCAGAACCTGGAGCGCCACATCTCCACTCTCAACATCCAGCTC 4194  
 Qy 1381 SerAspSerLysLysLysLeuGlnAspPheAlaSerThrValGluAlaLeuGluGluGly 1400  
 |||||  
 Db 4195 TCCGACTCGAAGAAGAAGCTGCAGGACTTTGCCAGCACCGTGAAGCTCTGGAAGAGGGG 4254  
 Qy 1401 LysLysArgPheGlnLysGluIleGluAsnLeuThrGlnGlnTyrGluGluLysAlaAla 1420  
 |||||  
 Db 4255 AAGAAGAGGTTCCAGAAGGAGATCGAGAACCTCACCCAGCAGTACGAGGAGAAGGCGGCC 4314  
 Qy 1421 AlaTyrAspLysLeuGluLysThrLysAsnArgLeuGlnGlnGluLeuAspAspLeuVal 1440  
 |||||  
 Db 4315 GCTTATGATAAACTGGAAAAGACCAAGAACAGGCTTCAGCAGGAGCTGGACGACCTGGTT 4374  
 Qy 1441 ValAspLeuAspAsnGlnArgGlnLeuValSerAsnLeuGluLysLysGlnArgLysPhe 1460  
 |||||  
 Db 4375 GTTGATTGGACAACCAGCGGCAACTCGTGTCCAACCTGGAAAAGAAGCAGAGGAAATTT 4434  
 Qy 1461 AspGlnLeuLeuAlaGluGluLysAsnIleSerSerLysTyrAlaAspGluArgAspArg 1480  
 |||||  
 Db 4435 GATCAGTTGTTAGCCGAGGAGAAAAACATCTCTTCCAAATACGCGGATGAGAGGGACAGA 4494  
 Qy 1481 AlaGluAlaGluAlaArgGluLysGluThrLysAlaLeuSerLeuAlaArgAlaLeuGlu 1500  
 |||||  
 Db 4495 GCTGAGGCAGAAGCCAGGGAGAAGGAAACCAAGGCCCTGTCCCTGGCTCGGGCCCTTGAA 4554  
 Qy 1501 GluAlaLeuGluAlaLysGluGluLeuGluArgThrAsnLysMetLeuLysAlaGluMet 1520  
 |||||  
 Db 4555 GAGGCCCTTGGAAGCCAAAGAGGAACCTCGAGCGGACCAACAAATGCTCAAAGCCGAAATG 4614  
 Qy 1521 GluAspLeuValSerSerLysAspAspValGlyLysAsnValHisGluLeuGluLysSer 1540  
 |||||  
 Db 4615 GAAGACCTGGTCAGCTCCAAGGATGACGTGGGCAAGAACGTCCATGAGCTGGAGAAGTCC 4674  
 Qy 1541 LysArgAlaLeuGluThrGlnMetGluGluMetLysThrGlnLeuGluGluLeuGluAsp 1560  
 |||||  
 Db 4675 AAGCGGGCCCTGGAGACCCAGATGGAGGAGATGAAGACGCAGCTGGAAGAGCTGGAGGAC 4734  
 Qy 1561 GluLeuGlnAlaThrGluAspAlaLysLeuArgLeuGluValAsnMetGlnAlaLeuLys 1580  
 |||||  
 Db 4735 GAGCTGCAAGCCACGAGGACGCCAAACTGCGGCTGGAAGTCAACATGCAGGCGCTCAAG 4794  
 Qy 1581 GlyGlnPheGluArgAspLeuGlnAlaArgAspGluGlnAsnGluGluLysArgArgGln 1600  
 |||||  
 Db 4795 GGCCAGTTCGAAAGGGATCTCCAAGCCCGGACGAGCAGAATGAGGAGAAGAGGAGGCAA 4854  
 Qy 1601 LeuGlnArgGlnLeuHisGluTyrGluThrGluLeuGluAspGluArgLysGlnArgAla 1620  
 |||||  
 Db 4855 CTGCAGAGACAGCTTCACAGTATGAGACGGAACCTGGAAGACGAGCGAAAGCAACGTGCC 4914  
 Qy 1621 LeuAlaAlaAlaAlaLysLysLysLeuGluGlyAspLeuLysAspLeuGluLeuGlnAla 1640  
 |||||  
 Db 4915 CTGGCAGCTGCAGCAAAGAAGAAGCTGGAAGGGGACCTGAAAGACCTGGAGCTTCAGGCC 4974  
 Qy 1641 AspSerAlaIleLysGlyArgGluGluAlaIleLysGlnLeuArgLysLeuGlnAlaGln 1660  
 |||||  
 Db 4975 GACTCTGCCATCAAGGGGAGGAGGAAGCCATCAAGCAGCTACGCAAACCTGCAGGCTCAG 5034  
 Qy 1661 MetLysAspPheGlnArgGluLeuGluAspAlaArgAlaSerArgAspGluIlePheAla 1680  
 |||||  
 Db 5035 ATGAAGGACTTTCAAAGAGAGCTGGAAGATGCCCGTGCTCCAGAGATGAGATCTTTGCC 5094  
 Qy 1681 ThrAlaLysGluAsnGluLysLysAlaLysSerLeuGluAlaAspLeuMetGlnLeuGln 1700

Db	5095		ACAGCCAAAGAGAATGAGAAGAAAGCCAAGAGCTTGAAGCAGACCTCATGCAGCTACAA	5154
Qy	1701		GluAspLeuAlaAlaAlaGluArgAlaArgLysGlnAlaAspLeuGluLysGluGluLeu	1720
Db	5155		GAGGACCTCGCCGCCGCTGAGAGGGCTCGAAACAAGCGACCTCGAGAAGGAGGAAGCTG	5214
Qy	1721		AlaGluGluLeuAlaSerSerLeuSerGlyArgAsnAlaLeuGlnAspGluLysArgArg	1740
Db	5215		GCAGAGGAGCTGGCCAGTAGCTGTCTGGGAAGGAACGCCTCCAGGACGAGAAGCGCCGC	5274
Qy	1741		LeuGluAlaArgIleAlaGlnLeuGluGluGluLeuGluGluGluGlnGlyAsnMetGlu	1760
Db	5275		CTGGAGGCCCGGATCGCCAGCTGGAGGAGGAGCTGGAGGAGGAGCAGGGCAACATGGAG	5334
Qy	1761		AlaMetSerAspArgValArgLysAlaThrGlnGlnAlaGluGlnLeuSerAsnGluLeu	1780
Db	5335		GCCATGAGCGACCGGGTCCGCAAAGCCACACAGCAGGCCGAGCAGCTCAGCAACGAGCTG	5394
Qy	1781		AlaThrGluArgSerThrAlaGlnLysAsnGluSerAlaArgGlnGlnLeuGluArgGln	1800
Db	5395		GCCACAGAGCGCAGCACGGCCAGAGAATGAGAGTGCCCGGCAGCAGCTCGAGCGGCAG	5454
Qy	1801		AsnLysGluLeuArgSerLysLeuHisGluMetGluGlyAlaValLysSerLysPheLys	1820
Db	5455		AACAAGGAGCTCCGGAGCAAGCTCCACGAGATGGAGGGGGCCGTCAAGTCCAAGTTCAA	5514
Qy	1821		SerThrIleAlaAlaLeuGluAlaLysIleAlaGlnLeuGluGluGlnValGluGlnGlu	1840
Db	5515		TCCACCATCGCGGCGCTGGAGGCCAAGATTGCACAGCTAGAGGAGCAGGTCGAGCAGGAG	5574
Qy	1841		AlaArgGluLysGlnAlaAlaThrLysSerLeuLysGlnLysAspLysLysLeuLysGlu	1860
Db	5575		GCCAGAGAGAAACAGGCGGCCACCAAGTCGCTGAAGCAGAAAGACAAGAAGCTGAAGGAA	5634
Qy	1861		IleLeuLeuGlnValGluAspGluArgLysMetAlaGluGlnTyrLysGluGlnAlaGlu	1880
Db	5635		ATCTTGCTGCAGGTGGAGGACGAGCGCAAGATGGCCGAGCAGTACAAGGAGCAGGCAGAG	5694
Qy	1881		LysGlyAsnAlaArgValLysGlnLeuLysArgGlnLeuGluGluAlaGluGluGluSer	1900
Db	5695		AAAGGCAATGCCAGGGTCAAGCAGCTCAAGAGGCAGCTGGAGGAGGCAGAGGAGGAGTCC	5754
Qy	1901		GlnArgIleAsnAlaAsnArgArgLysLeuGlnArgGluLeuAspGluAlaThrGluSer	1920
Db	5755		CAGCGCATCAACGCCAACCGCAGGAAGCTGCAGCGGGAGCTGGATGAGGCCACGGAGAGC	5814
Qy	1921		AsnGluAlaMetGlyArgGluValAsnAlaLeuLysSerLysLeuArg	1936
Db	5815		AACGAGGCCATGGGCCGCGAGGTGAACGCACTCAAGAGCAAGCTCAGG	5862

# Sequence Alignment

## GenPept Accession Number P35749

MYHB\_HUMAN

ID MYHB\_HUMAN STANDARD; PRT; 1972 AA.  
AC P35749; O00396; O94944; P78422;  
DT 01-JUN-1994 (Rel. 29, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 15-MAR-2004 (Rel. 43, Last annotation update)  
DE Myosin heavy chain, smooth muscle isoform (SMMHC).  
GN MYH11 OR KIAA0866.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=99425270; PubMed=10493829;  
RA Loftus B.J., Kim U.-J., Sneddon V.P., Kalush F., Brandon R.,  
RA Fuhrmann J., Mason T., Crosby M.L., Barnstead M., Cronin L.,  
RA Deslattes Mays A., Cao Y., Xu R.X., Kang H.-L., Mitchell S.,  
RA Eichler E.E., Harris P.C., Venter J.C., Adams M.D.;  
RT "Genome duplications and other features in 12 Mb of DNA sequence from  
RT human chromosome 16p and 16q.";  
RL Genomics 60:295-308(1999).  
RN [2]  
RP SEQUENCE FROM N.A., AND REVISIONS TO 1263-1266.  
RC TISSUE=Brain;  
RA Nagase T., Kikuno R., Yamakawa H., Ohara O.;  
RL Submitted (JAN-2003) to the EMBL/GenBank/DDBJ databases.  
RN [3]  
RP SEQUENCE OF 1-1266 FROM N.A.  
RC TISSUE=Brain;  
RX MEDLINE=99156230; PubMed=10048485;  
RA Nagase T., Ishikawa K.-I., Suyama M., Kikuno R., Hirose M.,  
RA Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.;  
RT "Prediction of the coding sequences of unidentified human genes. XII.  
RT The complete sequences of 100 new cDNA clones from brain which code  
RT for large proteins in vitro.";  
RL DNA Res. 5:355-364(1998).  
RN [4]  
RP SEQUENCE OF 885-1972 FROM N.A.  
RX MEDLINE=93263189; PubMed=7684189;  
RA Matsuoka R., Yoshida M.C., Furutani Y., Imamura S., Kanda N.,  
RA Yanagisawa M., Masaki T., Takao A.;  
RT "Human smooth muscle myosin heavy chain gene mapped to chromosomal  
RT region 16q12.";  
RL Am. J. Med. Genet. 46:61-67(1993).  
RN [5]  
RP SEQUENCE OF 1093-1972 FROM N.A.  
RC TISSUE=Hippocampus;  
RA Okajima K.;  
RL Submitted (NOV-1992) to the EMBL/GenBank/DDBJ databases.  
CC -!- FUNCTION: Muscle contraction.  
CC -!- SUBUNIT: Muscle myosin is a hexameric protein that consists of 2  
CC heavy chain subunits (MHC), 2 alkali light chain subunits (MLC)  
CC and 2 regulatory light chain subunits (MLC-2).  
CC -!- SUBCELLULAR LOCATION: Thick filaments of the myofibrils.  
CC -!- TISSUE SPECIFICITY: Smooth muscle; expressed in the umbilical  
CC artery, bladder, esophagus and trachea.  
CC -!- DOMAIN: The rodlike tail sequence is highly repetitive, showing  
CC cycles of a 28-residue repeat pattern composed of 4 heptapeptides,  
CC characteristic for alpha-helical coiled coils.  
CC -!- DISEASE: A chromosomal rearrangement, known as pericentric  
CC inversion Inv(16)(p13;q22), Produces a fusion protein that  
CC consists of the 165 N-terminal residues of CBF-beta (PEPB2) with  
CC the tail region of MYH11. This rearrangement is associated with  
CC acute myeloid leukemia of M4EO subtype.  
CC -!- MISCELLANEOUS: Each myosin heavy chain can be split into 1 light  
CC meromyosin (LMM) and 1 heavy meromyosin (HMM). It can later be  
CC split further into 2 globular subfragments (S1) and 1 rod-shaped  
CC subfragment (S2).  
CC -!- SIMILARITY: Contains 1 myosin-like globular head domain.



CC -!- SIMILARITY: Contains 1 IQ domain.  
CC -!- DATABASE: NAME=Atlas Genet. Cytogenet. Oncol. Haematol.;  
CC WWW="http://www.infobiogen.fr/services/chromcancer/Genes/MYH11ID43.html".  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
CC or send an email to license@isb-sib.ch).  
CC -----  
DR EMBL; AF001548; AAC31665.1; -.  
DR EMBL; U91323; AAC35212.1; -.  
DR EMBL; AB020673; BAA74889.2; ALT\_INIT.  
DR EMBL; D10667; -; NOT\_ANNOTATED\_CDS.  
DR EMBL; X69292; CAA49154.1; -.  
DR HSSP; P10587; 1BR2.  
DR Genew; HGNC:7569; MYH11.  
DR MIM; 160745; -.  
DR GO; GO:0005859; C:muscle myosin; TAS.  
DR InterPro; IPR000048; IQ\_region.  
DR InterPro; IPR001609; myosin\_head.  
DR InterPro; IPR004009; Myosin\_N.  
DR InterPro; IPR002928; Myosin\_tail.  
DR InterPro; IPR002017; Spectrin.  
DR Pfam; PF00612; IQ; 1.  
DR Pfam; PF00063; myosin\_head; 1.  
DR Pfam; PF02736; Myosin\_N; 1.  
DR Pfam; PF01576; Myosin\_tail; 1.  
DR PRINTS; PR00193; MYOSINHEAVY.  
DR ProDom; PD000355; myosin\_head; 1.  
DR SMART; SM00015; IQ; 1.  
DR SMART; SM00242; MYSc; 1.  
DR PROSITE; PS50096; IQ; 1.  
KW Myosin; Muscle protein; Coiled coil; Thick filament; Actin-binding;  
KW Calmodulin-binding; ATP-binding; Methylation; Multigene family;  
KW Proto-oncogene; Chromosomal translocation.  
FT DOMAIN 1 785 MYOSIN HEAD-LIKE.  
FT DOMAIN 786 815 IQ.  
FT DOMAIN 844 1934 COILED COIL (POTENTIAL).  
FT DOMAIN 1935 1972 CARBOXYL-TERMINAL.  
FT NP\_BIND 178 185 ATP (POTENTIAL).  
FT DOMAIN 661 683 ACTIN-BINDING (BY SIMILARITY).  
FT DOMAIN 762 776 ACTIN-BINDING (BY SIMILARITY).  
FT MOD\_RES 129 129 METHYLATION (TRI-) (POTENTIAL).  
FT CONFLICT 887 889 EEK -> NSE (IN REF. 4).  
FT CONFLICT 1558 1558 T -> S (IN REF. 4).  
FT CONFLICT 1610 1611 KQ -> NE (IN REF. 4).  
FT CONFLICT 1786 1786 A -> S (IN REF. 5).  
FT CONFLICT 1958 1958 T -> L (IN REF. 4).  
SQ SEQUENCE 1972 AA; 227338 MW; 67665BB2AECE1277 CRC64;

Query Match 98.9%; Score 9602.5; DB 1; Length 1972;  
Best Local Similarity 99.6%; Pred. No. 3e-279;  
Matches 1929; Conservative 0; Mismatches 0; Indels 7; Gaps 1;

```
QY      1 MAQKGQLSDDEKFLFVDKNFINSPVAQADWAAKRLVWVPSEKQGFEAASIKEEKGDEVVV 60
      |||
Db      1 MAQKGQLSDDEKFLFVDKNFINSPVAQADWAAKRLVWVPSEKQGFEAASIKEEKGDEVVV 60

QY     61 ELVENGKKVTVGKDDIQMNPFPKFSKVEDMAELTCLNEASVLHNLRLERYFSGLIYTYSG 120
      |||
Db     61 ELVENGKKVTVGKDDIQMNPFPKFSKVEDMAELTCLNEASVLHNLRLERYFSGLIYTYSG 120

QY    121 FCVVVNPNYKHLPIYSEKIVDMYKGGKRHEMPPHIYAIADTAYRSMQLDREDQSILCTGES 180
      |||
Db    121 FCVVVNPNYKHLPIYSEKIVDMYKGGKRHEMPPHIYAIADTAYRSMQLDREDQSILCTGES 180

QY    181 GAGKTENTKKVIQYLAVVASSHKGKDKTSTITQGPSFAYGELEKQLLQANPILEAFGNAKT 240
      |||
```

Db 181 GAGKTENTKKVIQYLAVVASSHKGKDKTSIT-----GELEKQLLQANPILEAFGNAKT 233

Qy 241 VKNDNSSRFGKFIRINFDVTGYIVGANIETYLLEKSRAIRQARDERTFHIFYMIAGAKE 300  
|||||

Db 234 VKNDNSSRFGKFIRINFDVTGYIVGANIETYLLEKSRAIRQARDERTFHIFYMIAGAKE 293  
|||||

Qy 301 KMRSDDLLEGFNNYTFLSNGFVPIPAAQDDEMFOETVEAMAIMGFSEEEQLSILKVSSV 360  
|||||

Db 294 KMRSDDLLEGFNNYTFLSNGFVPIPAAQDDEMFOETVEAMAIMGFSEEEQLSILKVSSV 353  
|||||

Qy 361 LQLGNIVFKKERNTDQASMPDNTAAQKVCHLMGINVTDFTRSILTPRIKVG RDVVQKAQT 420  
|||||

Db 354 LQLGNIVFKKERNTDQASMPDNTAAQKVCHLMGINVTDFTRSILTPRIKVG RDVVQKAQT 413  
|||||

Qy 421 KEQADFAVEALAKATYERLFRWILTRVNKALDKTHRQASFLGILDIAGFEIFEVNSFEQ 480  
|||||

Db 414 KEQADFAVEALAKATYERLFRWILTRVNKALDKTHRQASFLGILDIAGFEIFEVNSFEQ 473  
|||||

Qy 481 LCINYTNEKLQQLFNHTMFILEQEYQREGIEWNFIDFGLDLQPCIELIERPNNPPGVLA 540  
|||||

Db 474 LCINYTNEKLQQLFNHTMFILEQEYQREGIEWNFIDFGLDLQPCIELIERPNNPPGVLA 533  
|||||

Qy 541 LLDEECWFPKATDKSFVEKLCTEQGSHPKFQPKQLKDKTEFSIIHYAGKVDYNASAWLT 600  
|||||

Db 534 LLDEECWFPKATDKSFVEKLCTEQGSHPKFQPKQLKDKTEFSIIHYAGKVDYNASAWLT 593  
|||||

Qy 601 KNMDPLNDNVTSLLNASSDKFVADLWKDVDRIVGLDQMAKMTESLPSASKTKKGMFRTV 660  
|||||

Db 594 KNMDPLNDNVTSLLNASSDKFVADLWKDVDRIVGLDQMAKMTESLPSASKTKKGMFRTV 653  
|||||

Qy 661 GQLYKEQLGKLMTTLRNTTPNFVRCIIPNHEKRSGLDAFLVLEQLRCNGVLEGIRICRQ 720  
|||||

Db 654 GQLYKEQLGKLMTTLRNTTPNFVRCIIPNHEKRSGLDAFLVLEQLRCNGVLEGIRICRQ 713  
|||||

Qy 721 GFPNRIVFQEFRQRYEILAANAIPKGFMDGKQACILMIKALELDPNLYRIGQSKIFFRTG 780  
|||||

Db 714 GFPNRIVFQEFRQRYEILAANAIPKGFMDGKQACILMIKALELDPNLYRIGQSKIFFRTG 773  
|||||

Qy 781 VLAHLEERDLKITDVIMAFQAMCRGYLARKAFARQQQLTAMKVIQRNCAAYLKLRNWQ 840  
|||||

Db 774 VLAHLEERDLKITDVIMAFQAMCRGYLARKAFARQQQLTAMKVIQRNCAAYLKLRNWQ 833  
|||||

Qy 841 WWRLFTKVKPLLQVTRQEEEMQAKDELQKTKERQQKAENELKELEQKHSQLTEKNLLQ 900  
|||||

Db 834 WWRLFTKVKPLLQVTRQEEEMQAKDELQKTKERQQKAENELKELEQKHSQLTEKNLLQ 893  
|||||

Qy 901 EQLQAETELYAEAEEMRVRLAAKKQLEELIHEMEARLEEEEDRGQQQAERKKMAQQML 960  
|||||

Db 894 EQLQAETELYAEAEEMRVRLAAKKQLEELIHEMEARLEEEEDRGQQQAERKKMAQQML 953  
|||||

Qy 961 DLEEQLBEEEAARQKLQLEKVTAEAIIKKLEDEILVMDDQNNKLSKERKLEERISDLTT 1020  
|||||

Db 954 DLEEQLBEEEAARQKLQLEKVTAEAIIKKLEDEILVMDDQNNKLSKERKLEERISDLTT 1013  
|||||

Qy 1021 NLAEELKAKNLTCLKNKHESMISELEVRLKKEEKSQLEKLKRKLEGDASDFHEQIAD 1080  
|||||

Db 1014 NLAEELKAKNLTCLKNKHESMISELEVRLKKEEKSQLEKLKRKLEGDASDFHEQIAD 1073  
|||||

Qy 1081 LQAQIAELKMLAKKEEELQAALARLDDEIAQKNNALKKIRELEGHISDLQEDLDSERAA 1140  
|||||

Db 1074 LQAQIAELKMLAKKEEELQAALARLDDEIAQKNNALKKIRELEGHISDLQEDLDSERAA 1133  
|||||

Qy 1141 RNKAQKQKRDLGEELEALKTELEDTLSTATQQLRAKREQEVTVLKKALDEETRSHEAQ 1200  
|||||

Db 1134 RNKAQKQKRDLGEELEALKTELEDTLSTATQQLRAKREQEVTVLKKALDEETRSHEAQ 1193  
|||||

Qy 1201 VQEMRQKHAQAVEELTEQLEQFKRANKLDKNKQTLKENADLAGELRVLGQAKQVEVHK 1260  
|||||

Db 1194 VQEMRQKHAQAVEELTEQLEQFKRANKLDKNKQTLKENADLAGELRVLGQAKQVEVHK 1253  
|||||

Qy 1261 KKKLEAQVQELQSKCSDGERARAEIENDKVHKLQNEVESVTGMLNEAEGKAIKLAQDVASL 1320

Db	1254	KKKLEAQVQELQSKCSDGERARAELNDKVHKLQNEVESVTGMLNEAEGKAIKLAKDVASL	1313
Qy	1321	SSQLQDTQELLQEETRQKLNVTSTKLRLQLEEBERNLQDQLDEEMEAKQNLERHISTLNIQL	1380
Db	1314	SSQLQDTQELLQEETRQKLNVTSTKLRLQLEEBERNLQDQLDEEMEAKQNLERHISTLNIQL	1373
Qy	1381	SDSKKKLQDFASTVEALEEGKKRFQKEIENLTQQYEEKAAAYDKLEKTKNRLQQELDDL	1440
Db	1374	SDSKKKLQDFASTVEALEEGKKRFQKEIENLTQQYEEKAAAYDKLEKTKNRLQQELDDL	1433
Qy	1441	VDLDNQRLVSNLEKKQRKFDQLLAEKNISSKYADERDRAEAEAREKETKALSLARALE	1500
Db	1434	VDLDNQRLVSNLEKKQRKFDQLLAEKNISSKYADERDRAEAEAREKETKALSLARALE	1493
Qy	1501	EALAEKEELERTNKMLKAEMEDLVSSKDDVGKNVHELEKSKRALETQMEEMKTQLEELED	1560
Db	1494	EALAEKEELERTNKMLKAEMEDLVSSKDDVGKNVHELEKSKRALETQMEEMKTQLEELED	1553
Qy	1561	ELQATEDAKLRLEVNMQALKGGQFERDLQARDEQNEEKRRQLQRLHEYETELEDERKQRA	1620
Db	1554	ELQATEDAKLRLEVNMQALKGGQFERDLQARDEQNEEKRRQLQRLHEYETELEDERKQRA	1613
Qy	1621	LAAAAKKKLEGDLKDLELQADSAIKGREEAIKQLRKLQAQMKDFQRELEDARASRDEIFA	1680
Db	1614	LAAAAKKKLEGDLKDLELQADSAIKGREEAIKQLRKLQAQMKDFQRELEDARASRDEIFA	1673
Qy	1681	TAKENEEKAKSLEADLMQLQEDLAAERARKQADLEKEELAEELASSLSGRNALQDEKRR	1740
Db	1674	TAKENEEKAKSLEADLMQLQEDLAAERARKQADLEKEELAEELASSLSGRNALQDEKRR	1733
Qy	1741	LEARIAQLEEELEEEQGNMEAMSDRVKATQQAEQLSNELATERSTAQKNESARQQLERQ	1800
Db	1734	LEARIAQLEEELEEEQGNMEAMSDRVKATQQAEQLSNELATERSTAQKNESARQQLERQ	1793
Qy	1801	NKELRSKLHEMEGAVKSKFKSTIAALEAKIAQLEEQVEQAEAREKQAATKSLKQDKKKLKE	1860
Db	1794	NKELRSKLHEMEGAVKSKFKSTIAALEAKIAQLEEQVEQAEAREKQAATKSLKQDKKKLKE	1853
Qy	1861	ILLQVEDERKMAEQYKEQAEKGNARVKQLKRQLEEAEEESQRINANRRKLQRELDEATES	1920
Db	1854	ILLQVEDERKMAEQYKEQAEKGNARVKQLKRQLEEAEEESQRINANRRKLQRELDEATES	1913
Qy	1921	NEAMGREVNALKSKLR	1936
Db	1914	NEAMGREVNALKSKLR	1929